

Philippe N. Johnson-Laird, Monica Bucciarelli

A Reply to Ratti and Brigaglia and Celano's Comments on Deontics: Meaning, Reasoning, and Emotion

(doi: 10.1436/92964)

Materiali per una storia della cultura giuridica (ISSN 1120-9607)

Fascicolo 1, giugno 2019

Ente di afferenza:

Universitgli studi di Genova (unige)

Copyright © by Società editrice il Mulino, Bologna. Tutti i diritti sono riservati.

Per altre informazioni si veda <https://www.rivisteweb.it>

Licenza d'uso

L'articolo è messo a disposizione dell'utente in licenza per uso esclusivamente privato e personale, senza scopo di lucro e senza fini direttamente o indirettamente commerciali. Salvo quanto espressamente previsto dalla licenza d'uso Rivisteweb, è fatto divieto di riprodurre, trasmettere, distribuire o altrimenti utilizzare l'articolo, per qualsiasi scopo o fine. Tutti i diritti sono riservati.

A REPLY TO RATTI AND BRIGAGLIA AND CELANO'S COMMENTS ON *DEONTICS: MEANING, REASONING, AND EMOTION*

by Philip N. Johnson-Laird and Monica Bucciarelli

This article replies to the criticisms of Ratti, and of Brigaglia and Celano, of the model theory of deontics. It argues that the theory is consistent and that logic is an implausible basis for human reasoning.

Keywords: Model theory, Deontics, Reasoning.

Foreword

Professors Ratti, and Brigaglia and Celano have made insightful and provocative comments on our model theory of deontics. They have more issues and queries than we can satisfy here, in part because we don't know the answers and in part because a paper dealing with all their points would tax readers too much. So, in what follows, we deal with some of their issues, and pass by others in silence. In any case, we thank these scholars for all their thoughts. They are an invaluable guide to how we might improve the model theory or its presentation, or both.

Here's a sketch of the model theory of deontics – to save you from having to read anything else. No simple way exists to pick out moral propositions from the rest of deontic claims, such as those about social conventions, games, and good manners. And the foundations of moral judgments are neither complete nor consistent. That is because morality depends on beliefs, and to establish their consistency is computationally intractable – to check that 100 beliefs could all be true might take longer than the likely survival of the solar system, because there could be 2^{100} cases to assess. Incompleteness implies that you should have difficulty in making certain judg-

Monica Bucciarelli. Dipartimento di Psicologia and Center for Logic, Language and Cognition, Università di Torino, Via Po 14, 10123 Torino, monica.bucciarelli@unito.it

Philip N. Johnson-Laird, Stuart Professor of Psychology, Emeritus, Princeton University, Visiting scholar, Department of Psychology, New York University, phil@princeton.edu

ments. A film star pays someone to keep the dirt about him from the public – it could be blackmail or an exercise in public relations. So, why is the former a crime and the latter legal? Is it right or wrong for a penniless mother to steal food to feed her starving children? It can be hard to say. According to the model theory, all such judgments depend on reasoning from models of the world. *Having children obligates you to earn a living* is a “weak” obligation in that there can be reasons other than having children for why you ought to earn a living. In contrast, *having a paid job obligates you to pay income tax* is a “strong” obligation in that such a job is the only reason for you to pay income tax. Deontic matters elicit emotional responses, but they derive from cognitions too crude to distinguish between the two sorts of obligation, and they alone cannot suffice to assess that a moral judgment is correct. Emotions antedate human beings by 400 million years or so. And the mammals in which they evolved have rudimentary powers of reasoning, but as far as we can tell they do not dabble in deontics. Factual and deontic assertions vary in believability, and both can elicit emotions. But, only deontic claims yield a striking correlation. People are Epicurean in judgment. They love deontic propositions they believe, and hate those they disbelieve. The mere substitution of the word *ought* for *is* in an assertion switches on this relation between belief and feeling¹.

1. *On Ratti*

Ratti made many criticisms of the model theory.

1.1. He argues that it is inconsistent, and necessarily so. It postulates that obligation, prohibition, and permissibility are interdefinable, and so people understand that no essential difference exists among these three claims:

It is permissible to smoke.
It is not obligatory not to smoke.
It is not prohibited to smoke.

¹ M. Bucciarelli, P.N. Johnson-Laird, *Emotions and Beliefs about Morality Can Change One Another*, forthcoming.

These equivalents are common to most modal logics². Ratti translates a version of the previous postulate into a modal sentential logic:

$$\text{Permissible } a \rightarrow \neg \text{Obligatory } \neg a$$

where “a” denotes an action or inaction, “ \rightarrow ” denotes the *material* conditional of sentential logic, which is true in any case except when its antecedent is true and its consequent is false, and “ \neg ” denotes sentential negation. He invokes the rational principle that an action cannot be both permissible and impermissible:

$$\neg (\text{Permissible } a \ \& \ \neg \text{Permissible } a)$$

(It is a special case of the general principle that a predicate either applies or doesn't apply to any action.) The model theory, as we mentioned, assumes that the foundations of moral judgments are incomplete and inconsistent. Which implies that an action can be both permissible and impermissible:

$$\text{Permissible } a \ \& \ \neg \text{Permissible } a$$

And this proposition, as Ratti points out, contradicts his rational principle, and so there is a huge flaw in the model theory: it is inconsistent, and necessarily so.

It's a striking claim, but misses its target. The model theory implies that *people* can be inconsistent. And Ratti's argument demonstrates how an inconsistency can arise, as in the case of the penniless mother. It's wrong for her to steal; and yet it's wrong for her not to steal. Even if you come down heavily on one side or the other, you should know that people can amend such dilemmas until they claim that they are unable to resolve them³. Yet, does it follow that a theory that predicts inconsistencies is itself inconsistent? Not at all. If it were so, no sensible theory could explain irrationality. Ratti has proved only that a set of propositions is inconsistent, not that the model theory is. His formulas in modal logic do not refer to any theory. He assigns their inconsistency to the model theory, but it assigns them to people.

² G.E. Hughes, M.J. Cresswell, *A New Introduction to Modal Logic*, London, Routledge, 1996.

³ M. Bucciarelli, S. Khemlani, P.N. Johnson-Laird, *The Psychology of Moral Reasoning*, in «Judgment and Decision Making», 3, 2008, pp. 121-139.

Whether individuals are *necessarily* inconsistent about moral principles, we doubt. In the “possible worlds” semantics of modal logics, if people were necessarily inconsistent then they would be inconsistent in any relevant alternative world to ours. Yet, it is easy to envisage a world in which people have no moral beliefs whatsoever – one such “world” might be found in Washington, DC right now – and so their beliefs about morality cannot be inconsistent, because they have none. Ergo, people can be inconsistent as a matter of fact, not necessity.

1.2. Ratti argues that the model theory’s distinction between strong and weak obligations is suspect. As he points out, strong obligations are akin to biconditionals, whereas weak obligations are akin to conditionals. Consider the following putative equivalence:

Being already married prohibits you from marrying your new lover.

and:

Not being already married permits you to marry your new lover.

He remarks of such equivalences that they are true only for particular deontic systems under particular circumstances. There are indeed cultures in which the first claim can be false and the second claim true. But, this fact casts no doubt on the existence of strong and weak obligations.

1.3. Logic is a supreme intellectual invention: no logic; no theory of computability; no computers. But, Ratti’s recourse to logic merits a caution. It is not a secure guide to good reasoning in daily life, for several reasons:

– In logic, infinitely many valid conclusions follow from any set of premises. In life, people often assert that nothing follows, as they are likely to do given the premises: *Sneezing is permissible; Not sneezing is permissible.*

– In logic, any conclusion whatsoever follows from a contradiction. In life, people refrain from drawing conclusions from contradictions. Instead, they seek to resolve them⁴. It is this so-called “monotonic” feature of logic, which we suspect Ratti has in mind when he mentions towards the end of his critique that a common

⁴ See P.N. Johnson-Laird, V. Girotto, P. Legrenzi, *Reasoning from Inconsistency to Consistency*, in «Psychological Review», 111, 2004, pp. 640-661.

view holds that any inconsistent system is bound to be complete. True in logic, but false in daily life. We are all likely to harbor inconsistencies: we just haven't discovered them. When we write down our thoughts, however, they often become apparent. Frege had a view about sets that led to a contradiction, as Russell discovered in 1901. The English author Jane Austen has a pug that undergoes an abrupt sex change during the course of her novel, *Mansfield Park*. Even Emma Bovary's eyes changed color. And there are whole websites devoted to inconsistencies in the Bible, the US constitution, and case law.

– In logic, the material conditional delivers sound deductions that are absurd, e.g.:

It is not the case that if the Christian God exists then atheism is correct.

Therefore, the Christian God exists.

The premise is true, because it negates an obvious falsehood, namely: If the Christian God exists then atheism is correct. The material conditional, as we pointed out above, is false only in the case that its antecedent is true and its consequent is false. So, its negation in the proof is equivalent to the conjunction:

The Christian God exists and atheism is not correct.

From which it follows validly:

The Christian God exists.

In sum, standard logic is not an infallible guide to the inferences of daily life. What's worse for Ratti's analysis is that a countable infinity of modal sentential logics exist⁵. Their axioms differ, and so do their semantics. It may matter which modal logic he has in mind.

In fact, Ratti appears to rely on von Wright's deontic logic. But, consider an inference that we owe to Kamp⁶:

You may go to the beach or you may go to the cinema.

Therefore, you may go to the beach.

⁵ G.E. Hughes, M.J. Cresswell, *A New Introduction to Modal Logic*, cit.

⁶ H. Kamp, *Free Choice Permission*, in «Proceedings of the Aristotelian Society, New Series», 74, 1973-74, pp. 57-74.

Such inferences are commonplace in daily life. But, no deontic logic – including von Wright’s (as he knew) – validates them, or the truth of:

Permissible (a or b or both) → Permissible a.

Indeed, no modal logic allows the inference:

The flaw is in the software or in the hardware, or both.
Therefore, it is possible that the flaw is in the software.

Yet, such inferences are also commonplace in daily life⁷. The moral for students of human reasoning is simple: Beware logic!

1.4. Ratti takes issue with the model theory’s predictions of illusory inferences about morality. In fact, the theory predicts such systematic and compelling fallacies in all domains of reasoning, and experiments have corroborated their occurrence⁸. A bad habit among critics of the theory is to take issue with a single illusion, propose an alternative explanation for it, and blithely ignore the rest⁹. Ratti takes issue with this example:

You are permitted to carry out only one of the following actions:
Action 1: take the apple or the orange, or both
Action 2: take the pear or the orange, or both
Are you permitted to take the orange?

Most people respond, “Yes”. But, to take the orange is to carry out both actions 1 and 2, which is not permitted. Ratti points out that “action” is ambiguous: it could refer to a single act, such as taking the orange, or to the disjunction of two actions as “action 1” does. We doubt whether the ambiguity perplexes people, because they cope well with control problems. One had the same premises as those above, but instead posed the question:

Are you permitted to take the pear?

⁷ T. Hinterecker, M. Knauff, P.N. Johnson-Laird, *Modality, Probability, and Mental Models*, in «Journal of Experimental Psychology: Learning, Memory, and Cognition», 42, 2016, pp. 1606-1620.

⁸ S.S. Khemlani, P.N. Johnson-Laird, *Illusions in Reasoning*, in «Minds and Machines», 27, 2017, pp. 11-35.

⁹ B. Geurts, *Entertaining Alternatives: Disjunctions as Modals*, in «Natural Language Semantics», 13, 2005, pp. 383-410.

Over 90% of the participants replied, “Yes”, which is correct. Moreover, comparable illusions occur without any need to refer to actions or norms. Here’s the simplest illusory inference that we know, which concerns two exclusive disjunctions, i.e., their two clauses cannot both be true:

Either the cake is on the table or else the pie is.
Either the cake is on the table or else the pie isn’t.
Is it possible that the cake is on the table?

Most people say, “yes”. The model theory postulates they construct two *mental* models of the possibilities for what’s on the table:

The cake	
	The pie

Then they think about the two analogous possibilities for the second disjunction. So, they conclude that it is possible that the cake is on the table. *Fully explicit* models for what’s on the table represent false clauses too (using negation to do so). For the first premise, they are:

The cake	Not the pie
Not the cake	The pie

And for the second premise, they are:

The cake	The pie
Not the cake	Not the pie

No possibility is common to the two premises, and so their conjunction yields the null model (similar to the empty set), and so it does not follow that it is possible that the pie is on the table.

1.5. Ratti doubts whether all moral evaluations depend on reasoning (conscious or unconscious). You have reflexes that are automatically elicited by certain events: a puff of air in your eye, and you blink. Other animals have “fixed action patterns” in which a particular stimulus elicits an automatic sequence of behaviors. But, no evidence supports the automatic occurrence of moral evaluations of actions. Emotional reactions are rapid, and can occur prior to a moral evaluation. But, as we argued earlier, the cognitions that elicit them are primitive. To evaluate any action as morally good or

morally bad depends on comparing it to at least one deontic principle. The evaluation is at least akin to instantiation in logic: All theft is wrong; that is theft; ergo, it is wrong. It seems unlikely that the emotional system represents general propositions. The issue, however, is empirical. And when people think aloud as they pass judgment, their remarks support the model theory¹⁰.

1.6. Finally, Ratti is against the view that no difference exists between deontic reasoning and regular reasoning about deontic topics. Modus tollens, he says, is valid for descriptive conditionals, but not for the following sort of premises:

If you earn a salary then you are obligated to pay taxes.
You are not obligated to pay taxes.
Therefore, you don't earn a salary.

The inference seems valid, but Ratti says it is invalid, because it derives an “is” from an “ought”. In fact, nothing is wrong in doing so: Given that you ought to leave soon, it follows that you haven't left yet. It is the converse sort of inference deriving an “ought” from an “is” that isn't kosher. We therefore maintain that there is no special process of deontic reasoning. It is nothing more than regular reasoning about deontic topics, which can take into account that counterexamples to a normative principle don't refute it, but violate it.

2. On Brigaglia and Celano

Brigaglia and Celano accept some aspects of the model theory. But, they say:

2.1. It is no news that the foundation of deontic claims is permissibility and its negation. And they cite Allan Gibbard, the Utilitarian moral philosopher, as a proponent of the idea. Fair enough. Of course, granted the interdefinability of *permissible*, *obligation*, and *prohibition*, deontic claims could all be based instead on the concept of obligation, or in an Orwellian way on prohibition. Or, all three concepts could be treated as on a par. For us, the issue is empirical, and our experiments support the hypothesis that permissibility is the basic concept – they also support the hypothesis that possibil-

¹⁰ M. Bucciarelli, S. Khemlani, P.N. Johnson-Laird, *The Psychology of Moral Reasoning*, cit.

ity rather than necessity is foundational for non-deontic modalities¹¹. The novelty of the hypothesis, if any, is that it inspired its own empirical corroboration

2.2. For Brigaglia and Celano, the core question about normativity is also due to Gibbard: what kind of mental state is it to represent something as “permitted” (or as “correct”)? They list some such mental states: an erroneous belief in a non-existent property, a peculiar emotion, a qualitative state of a special kind, irreducible to beliefs or emotions. We write computer programs to implement our theories, and the analogous problem for a computer program is humdrum. It is a state that represents that an action (or not-action) is permitted. That’s all. The sense of “correctness” is more interesting. For some problems, you can have a sense that an answer is correct prior to your ability to articulate why it is correct – an observation that goes back to Bartlett¹², who used solutions to clues in crosswords as his example. Alas, sometimes the sense of “correctness” is wrong: in fact, you haven’t solved the problem. However, your deontic judgment that some action or inaction is correct should also elicit a corresponding emotion. You like what’s right and hate what’s wrong.

2.3. Brigaglia and Celano understand the distinction between mental models and fully explicit models. They ask a series of questions, which we can quote:

But what are fully explicit models? What is their status? Are they simply part of our psychological make-up, a set of rules deeply encoded in our mind, but only accessible with effort, and therefore normally neglected, or do they have an extra-psychological reality? In the former hypothesis, the rules of logic are indeed part of our psychological make-up (as for the mental logic model), but we usually take shortcuts (which explains our making systematic errors).

Let us try to answer them. You have a perceptual system that constructs models of the world. You may think you’re in direct contact with the world, but in fact a long causal chain occurs in which light bounces off surfaces, is focused on your retinae, and is processed in areas of your brain that yield a vivid kinematic representation of the scene. You can also imagine such scenes, constructing kinematic models¹³.

¹¹ See M. Ragni, P.N. Johnson-Laird, *Possibilities: A Theory of Naive Modal Reasoning*, 2019, under submission.

¹² F.C. Bartlett, *Thinking*, Cambridge, Cambridge University Press, 1958.

¹³ See M. Bucciarelli, R. Mackiewicz, S.S. Khemlani, P.N. Johnson-Laird, *Children’s Creation of Algorithms: Simulations and Gestures*, in «Journal of Cognitive Psychology», 28, 2016, pp. 297-318.

Other models derive from discourse. Descriptions tend to yield simple schematic representations, i.e., mental models representing possibilities. Their chief characteristic is that they focus on what is true. But, for simple tasks, you can construct fully explicit models, which also represent what is false. The contrast between the two sorts of model is illustrated in our earlier analysis of the problem of the cake. Models have an extra-psychological reality insofar as they control your behavior, which in turn can yield evidence corroborating their existence. You use models to reason: you can draw valid conclusions from them, which are *necessary* if they hold in all the fully explicit models of the premises, which are *possible* if they hold in at least one such model, and which are *probable* if they hold in most of a set of equiprobable models. Shortcuts may explain some errors, but in the model theory errors occur from a reliance on mental models rather than fully explicit ones, and from the failure to envisage all possible models of the premises. The biggest single source of human error is to overlook a possibility. As to the rules of logic, they play no part in the model theory, though some psychologists may still adhere to them as the basis of reasoning. No robust evidence corroborates their existence in the mind, except for individuals who have learned formal logic. In contrast, considerable evidence exists, including a whole variety of illusory inferences, to support the model theory.

Nowadays, most psychologists have abandoned logic as the basis for human reasoning. But, some of them suppose that the probability calculus provides such a foundation. We are skeptical, but it would take us far afield to explain why, other than to say: human beings don't know beans about probabilities. Two pilots appealed to one of us to resolve their disagreement about the probability that both engines of a twin-engined plane fail. The pilot who had flown F-16 jets (at more than 5 thousand kilometers per hour) said: double the probability of one engine failing. The glider pilot, who had run a company for the forensic analysis of software, said: halve it. Both of them were wrong. You square it if the two events are independent of one another. This anecdote bears out a robust experimental result¹⁴: unless you have studied the probability calculus, you won't know how to calculate the probability of a conjunction of two events. Perhaps it is no surprise that the British High Court banned the use of Bayes's theorem in evidence.

¹⁴ S. Khemlani, M. Lotstein, P.N. Johnson-Laird, *Naive Probability: Model-Based Estimates of Unique Events*, in «Cognitive Science», 39, 2015, pp. 1216-1258.